

We claim:

1. A method of enhancing audio renderings of non-audio data sources, comprising steps of:
  - detecting a nuance of a non-audio data source;
  - locating an audio cue corresponding to the detected nuance; and
  - associating the located audio cue with the detected nuance for playback to a listener.
2. The method according to Claim 1, further comprising the steps of:
  - creating an audio rendering of a non-audio segment of the non-audio data source, wherein the non-audio segment is associated with the nuance; and
  - mixing the associated audio cue with the audio rendering of the segment.
3. The method according to Claim 1, wherein the detecting step detects a plurality of nuances of the non-audio data source, the locating step locates audio cues for each of the detected nuances, and the associating step associates each of the located audio cues with the respective detected nuance, and further comprising the steps of:
  - creating an audio rendering of the non-audio data source; and
  - mixing the associated audio cues in with the audio rendering.
4. The method according to Claim 3, wherein the mixing step occurs while playing the audio rendering to the listener.

1 5. The method according to Claim 2 or Claim 3, wherein the non-audio data source is a text  
2 file and wherein the creating step further comprises processing the text file with a text-to-speech  
3 translator.

1 6. The method according to Claim 3, wherein at least one of the detected nuances is presence  
2 of a formatting tag.

1 7. The method according to Claim 3, wherein the non-audio data source is a text file and at  
2 least one of the detected nuances is a change in color of text in the text file.

1 8. The method according to Claim 1, wherein the non-audio data source is a text file and the  
2 detected nuance is a change in font of text in the text file.

1 9. The method according to Claim 1, wherein the non-audio data source is a text file and the  
2 detected nuance is presence of a keyword for the text file.

1 10. The method according to Claim 9, wherein the keyword is supplied by a creator of the text  
2 file.

1 11. The method according to Claim 9, wherein the keyword is programmatically detected by  
2 evaluating text in the text file.

1 12. The method according to Claim 3, wherein the non-audio data source is a text file and at  
2 least one of the detected nuances is presence of an emoticon in the text file.

1 13. The method according to Claim 1, wherein the detected nuance is a change of topic in the  
2 non-audio data source.

1 14. The method according to Claim 6, wherein the formatting tag is a new paragraph tag.

1 15. The method according to Claim 3, wherein at least one of the detected nuances is a degree  
2 of certainty in translation of the non-audio data source from another format.

1 16. The method according to Claim 15, wherein the detecting step detects at least two  
2 different degrees of certainty, and wherein the located audio cues comprise changes in a pitch of a  
3 voice used in the audio rendering for each of the different degrees of certainty.

1 17. The method according to Claim 15, wherein the detecting step detects at least two  
2 different degrees of certainty, and further comprising changing a pitch of the associated audio cue  
3 used by the mixing step for each of the different degrees of certainty.

1 18. The method according to Claim 15, wherein the detecting step detects at least two  
2 different degrees of certainty, and wherein the mixing step further comprises alternating between  
3 two of the located audio cues to audibly indicate the different degrees of certainty.

1 19. The method according to Claim 15, wherein the other format is an input audio data source  
2 and the non-audio data source is a text file, and the translation is an audio-to-text translation from  
3 the input audio data source to the text file, and wherein the degree of certainty reflects accuracy  
4 of the audio-to-text translation.

1 20. The method according to Claim 15, wherein the other format is an input audio data source  
2 and the non-audio data source is a text file, and the translation is an audio-to-text translation from  
3 the input audio data source to the text file, and wherein the degree of certainty reflects  
4 identification of a speaker who created the input audio data source.

1 21. The method according to Claim 15, wherein the other format is a source text file and the  
2 non-audio data source is an output text file, and the translation is a text-to-text translation from  
3 the source text file to the output text file, and wherein the degree of certainty reflects accuracy of  
4 the text-to-text translation.

1 22. The method according to Claim 21, wherein the source text file contains text in a first  
2 language and the output text file contains text in a second language.

1 23. The method according to Claim 3, wherein at least one of the detected nuances is an  
2 identification of a creator of the non-audio data source.

1 24. The method according to Claim 23, wherein the identification is used to locate stored  
2 preferences of the creator.

1 25. The method according to Claim 3, wherein the non-audio data source is an e-mail message  
2 and at least one of the detected nuances is an e-mail convention found in the e-mail message.

1 26. The method according to Claim 1, wherein the non-audio data source is text provided by a  
2 user.

1 27. The method according to Claim 26, wherein the text provided by the user is typed as  
2 command line input.

1 28. The method according to Claim 1, wherein the detected nuance is embedded within the  
2 non-audio file.

1 29. The method according to Claim 1, wherein the detected nuance comprises metadata  
2 associated with the non-audio file.

1 30. The method according to Claim 3, wherein the mixing step further comprises mixing in a  
2 streaming audio source for at least one of the located audio cues.

1 31. A method of enhancing audio renderings of data sources, comprising steps of:

transforming a first data source in a first format to a second data source in an audio format;  
associating one or more degrees of certainty with the second data source to reflect an accuracy of the transforming step;  
locating an audio cue that is correlated to each of the associated degrees of certainty; and  
associating the located audio cues with the second data source to convey the accuracy of the transforming step to a listener who will hear the audio format.

32. The method according to Claim 31, further comprising the step of audibly rendering the second data source to the listener along with the associated audio cues.

33. The method according to Claim 31, further comprising the step of storing the association of the located audio cues for subsequent audible rendering of the second data source to the listener along with the associated audio cues.

34. A method of enhancing audio renderings of non-audio data sources, comprising steps of:  
providing a stylesheet comprising rules and actions, wherein selected ones of the rules and actions pertain to audio cues to be used in an audio rendering;  
comparing the rules of the stylesheet to content of a non-audio data source; and  
upon detecting a match during the comparing step, applying the action associated with the matching rule, wherein for each action pertaining to audio cues, an audio cue is thereby associated with the non-audio data source for playing the audio rendering to a listener.

1 35. The method according to Claim 34, further comprising the step of playing the audio  
2 rendering of the non-audio data source to the listener.

1 36. The method according to Claim 34, wherein at least one of the selected rules and actions  
2 of the stylesheet is customized for the listener, and at least one of the audio cues associated with  
3 the non-audio data source by the applying step overrides another audio cue in order to customize  
4 the audio rendering for the listener.

1 37. The method according to Claim 35, wherein at least one of the audio cues associated with  
2 the non-audio data source by the applying step changes a pitch of a speaker's voice used in the  
3 playing step.

1 38. The method according to Claim 34, wherein at least one of the selected rules and actions  
2 of the stylesheet is customized for a creator of the non-audio data source, and at least one of the  
3 audio cues associated with the non-audio data source by the applying step overrides another audio  
4 cue in order to make the audio rendering speaker-specific.

1 39. The method according to Claim 34, wherein the stylesheet is an Extensible Stylesheet  
2 Language ("XSL") stylesheet.

1 40. The method according to Claim 35, wherein the stylesheet specifies preferences for  
2 language translation of the non-audio data source that may be performed prior to operation of the  
3 playing step.

1 41. A method of merchandising pre-recorded audio cues, further comprising steps of:  
2 receiving requests for selected ones of the pre-recorded audio cues for use as background  
3 sounds to be mixed with audibly rendered messages in order to provide enhanced contextual  
4 information to a listener of the audibly rendered messages; and  
5 providing the selected ones, in response to the step of receiving requests.

1 42. The method according to Claim 41, wherein the provided ones are used as an audio cue  
2 library.

1 43. A system for enhancing audio renderings of non-audio data sources, comprising:  
2 means for detecting one or more nuances of a non-audio data source;  
3 means for locating an audio cue corresponding to each of the detected nuances; and  
4 means for associating the located audio cues with their respective detected nuances for  
5 playback to a listener.

1 44. The system according to Claim 43, further comprising:  
2 means for creating an audio rendering of the non-audio data source, wherein the non-  
3 audio segment is associated with the nuance; and



4 means for mixing the associated audio cues in with the audio rendering while playing the  
5 audio rendering to the listener.

1 45. The system according to Claim 44, wherein the non-audio data source is a text file and  
2 wherein the means for creating further comprises means for processing the text file with a text-to-  
3 speech translator.

1 46. The system according to Claim 43, wherein at least one of the detected nuances is  
2 presence of a formatting tag.

1 47. The system according to Claim 43, wherein the non-audio data source is a text file and the  
2 detected nuance is a change in font of text in the text file.

1 48. The system according to Claim 43, wherein the non-audio data source is a text file and at  
2 least one of the detected nuances is presence of an emoticon in the text file.

1 49. The system according to Claim 43, wherein the detected nuance is a change of topic in the  
2 non-audio data source.

1 50. The system according to Claim 46, wherein the formatting tag is a new paragraph tag.

1 51. The system according to Claim 43, wherein at least one of the detected nuances is a  
2 degree of certainty in translation of the non-audio data source from another format.

1 52. The system according to Claim 51, wherein the means for detecting detects at least two  
2 different degrees of certainty, and wherein the located audio cues comprise changes in a pitch of a  
3 voice used in the audio rendering for each of the different degrees of certainty.

1 53. The system according to Claim 51, wherein the means for detecting detects at least two  
2 different degrees of certainty, and further comprising means for changing a pitch of the associated  
3 audio cue used by the means for mixing for each of the different degrees of certainty.

1 54. The system according to Claim 51, wherein the other format is an input audio data source  
2 and the non-audio data source is a text file, and the translation is an audio-to-text translation from  
3 the input audio data source to the text file, and wherein the degree of certainty reflects accuracy  
4 of the audio-to-text translation.

1 55. The system according to Claim 51, wherein the other format is an input audio data source  
2 and the non-audio data source is a text file, and the translation is an audio-to-text translation from  
3 the input audio data source to the text file, and wherein the degree of certainty reflects  
4 identification of a speaker who created the input audio data source.

1 56. The system according to Claim 51, wherein the other format is a source text file and the  
2 non-audio data source is an output text file, and the translation is a text-to-text translation from  
3 the source text file to the output text file, and wherein the degree of certainty reflects accuracy of  
4 the text-to-text translation.

1 57. The system according to Claim 43, wherein the non-audio data source is an e-mail  
2 message and at least one of the detected nuances is an e-mail convention found in the e-mail  
3 message.

1 58. The system according to Claim 43, wherein the non-audio data source is text provided by  
2 a user.

1 59. The system according to Claim 43, wherein the detected nuance is embedded within the  
2 non-audio file.

1 60. The system according to Claim 43, wherein the detected nuance comprises metadata  
2 associated with the non-audio file.

1 61. A system for enhancing audio renderings of data sources, comprising:  
2 means for transforming a first data source in a first format to a second data source in an  
3 audio format;

4 means for associating one or more degrees of certainty with the second data source to  
5 reflect an accuracy of the means for transforming;

6 means for locating an audio cue that is correlated to each of the associated degrees of  
7 certainty; and

8 means for associating the located audio cues with the second data source to convey the  
9 accuracy of the means for transforming to a listener who will hear the audio format.

1 62. The system according to Claim 61, further comprising means for audibly rendering the  
2 second data source to the listener along with the associated audio cues.

3 63. A system for enhancing audio renderings of non-audio data sources, comprising:

4 means for providing a stylesheet comprising rules and actions, wherein selected ones of  
5 the rules and actions pertain to audio cues to be used in an audio rendering;

6 means for comparing the rules of the stylesheet to content of a non-audio data source; and

7 means for applying the action associated with the matching rule, upon detecting a match  
8 during the comparing, wherein for each action pertaining to audio cues, an audio cue is thereby  
9 associated with the non-audio data source for playing the audio rendering to a listener.

1 64. The system according to Claim 62, further comprising means for playing the audio  
2 rendering of the non-audio data source to the listener.

1 65. The system according to Claim 63, wherein at least one of the selected rules and actions of  
2 the stylesheet is customized for the listener, and at least one of the audio cues associated with the  
3 non-audio data source by the means for applying overrides another audio cue in order to  
4 customize the audio rendering for the listener.

1 66. The system according to Claim 63, wherein at least one of the selected rules and actions of  
2 the stylesheet is customized for a creator of the non-audio data source, and at least one of the  
3 audio cues associated with the non-audio data source by the means for applying overrides another  
4 audio cue in order to make the audio rendering speaker-specific.

1 67. A computer program product for enhancing audio renderings of non-audio data sources,  
2 the computer program product embodied on one or more computer-readable media and  
3 comprising:

4 computer-readable program code means for detecting one or more nuances of a non-audio  
5 data source;

6 computer-readable program code means for locating an audio cue corresponding to each  
7 of the detected nuances; and

8 computer-readable program code means for associating the located audio cues with their  
9 respective detected nuances for playback to a listener.

1 68. The computer program product according to Claim 67, further comprising:

2 computer-readable program code means for creating an audio rendering of a non-audio  
3 segment of the non-audio data source, wherein the non-audio segment is associated with the  
4 nuance; and

5 computer-readable program code means for mixing the associated audio cue with the  
6 audio rendering of the segment.

1 69. The computer program product according to Claim 68, wherein the non-audio data source  
2 is a text file and wherein the computer-readable program code means for creating further  
3 comprises computer-readable program code means for processing the text file with a text-to-  
4 speech translator.

1 70. The computer program product according to Claim 67, wherein the non-audio data source  
2 is a text file and at least one of the detected nuances is a change in color of text in the text file.

1 71. The computer program product according to Claim 67, wherein the non-audio data source  
2 is a text file and the detected nuance is presence of a keyword for the text file.

1 72. The computer program product according to Claim 71, wherein the keyword is supplied  
2 by a creator of the text file.

1 73. The computer program product according to Claim 71, wherein the keyword is  
2 programmatically detected by evaluating text in the text file.

1 74. The computer program product according to Claim 67, wherein at least one of the  
2 detected nuances is a degree of certainty in translation of the non-audio data source from another  
3 format.

1 75. The computer program product according to Claim 74, wherein the computer-readable  
2 program code means for detecting detects at least two different degrees of certainty, and wherein  
3 the located audio cues comprise changes in a pitch of a voice used in the audio rendering for each  
4 of the different degrees of certainty.

1 76. The computer program product according to Claim 74, wherein the computer-readable  
2 program code means for detecting detects at least two different degrees of certainty, and further  
3 comprising changing a pitch of the associated audio cue used by the computer-readable program  
4 code means for mixing for each of the different degrees of certainty.

1 77. The computer program product according to Claim 74, wherein the other format is an  
2 input audio data source and the non-audio data source is a text file, and the translation is an  
3 audio-to-text translation from the input audio data source to the text file, and wherein the degree  
4 of certainty reflects accuracy of the audio-to-text translation.

1 78. The computer program product according to Claim 74, wherein the other format is an  
2 input audio data source and the non-audio data source is a text file, and the translation is an

3 audio-to-text translation from the input audio data source to the text file, and wherein the degree  
4 of certainty reflects identification of a speaker who created the input audio data source.

1 79. The computer program product according to Claim 74, wherein the other format is a  
2 source text file and the non-audio data source is an output text file, and the translation is a text-  
3 to-text translation from the source text file to the output text file, and wherein the degree of  
4 certainty reflects accuracy of the text-to-text translation.

1 80. The computer program product according to Claim 79, wherein the source text file  
2 contains text in a first language and the output text file contains text in a second language.

1 81. The computer program product according to Claim 67, wherein at least one of the  
2 detected nuances is an identification of a creator of the non-audio data source.

1 82. The computer program product according to Claim 81, wherein the identification is used  
2 to locate stored preferences of the creator.

1 83. The computer program product according to Claim 67, wherein the non-audio data source  
2 is an e-mail message.

1 84. The computer program product according to Claim 67, wherein the detected nuance is  
2 embedded within the non-audio file.



1 85. The computer program product according to Claim 67, wherein the detected nuance  
2 comprises metadata associated with the non-audio file.

1 86. A computer program product for enhancing audio renderings of data sources, the  
2 computer program product embodied on one or more computer-readable media and comprising:

3 computer-readable program code means for transforming a first data source in a first  
4 format to a second data source in an audio format;

5 computer-readable program code means for associating one or more degrees of certainty  
6 with the second data source to reflect an accuracy of the computer-readable program code means  
7 for transforming;

8 computer-readable program code means for locating an audio cue that is correlated to  
9 each of the associated degrees of certainty; and

10 computer-readable program code means for associating the located audio cues with the  
11 second data source to convey the accuracy of the computer-readable program code means for  
12 transforming to a listener who will hear the audio format.

1 87. The computer program product according to Claim 86, further comprising computer-  
2 readable program code means for audibly rendering the second data source to the listener along  
3 with the associated audio cues.

1 88. A computer program product for enhancing audio renderings of non-audio data sources,  
2 the computer program product embodied on one or more computer-readable media and  
3 comprising:

4 computer-readable program code means for comparing the rules of a stylesheet to content  
5 of a non-audio data source, wherein the stylesheet comprises rules and actions and wherein  
6 selected ones of the rules and actions pertain to audio cues to be used in an audio rendering; and

7 computer-readable program code means for applying the action associated with the  
8 matching rule, upon detecting a match during operation of the computer-readable program code  
9 means for comparing, wherein for each action pertaining to audio cues, an audio cue is thereby  
10 associated with the non-audio data source for playing the audio rendering to a listener.

11 89. The computer program product according to Claim 88, further comprising computer-  
12 readable program code means for playing the audio rendering of the non-audio data source to the  
13 listener.

14 90. The computer program product according to Claim 88, wherein at least one of the  
15 selected rules and actions of the stylesheet is customized for the listener, and at least one of the  
16 audio cues associated with the non-audio data source by the computer-readable program code  
17 means for applying overrides another audio cue in order to customize the audio rendering for the  
18 listener.

1 91. The computer program product according to Claim 88, wherein at least one of the  
2 selected rules and actions of the stylesheet is customized for a creator of the non-audio data  
3 source, and at least one of the audio cues associated with the non-audio data source by the  
4 computer-readable program code means for applying overrides another audio cue in order to  
5 make the audio rendering speaker-specific.

1 92. The computer program product according to Claim 89, wherein the stylesheet specifies  
2 preferences for language translation of the non-audio data source that may be performed prior to  
3 operation of the computer-readable program code means for playing.